

# FUNCIONALIDAD DE CUATRO MODELOS PARA LAS EVALUACIONES GENOTIPICAS EN SERIES DE EXPERIMENTOS

## FUNCTIONALITY OF FOUR MODELS FOR GENOTYPIC EVALUATIONS IN SERIES OF EXPERIMENTS

Jaime Sahagún Castellanos<sup>1</sup>

### RESUMEN

La calidad del análisis de la información producida en series de experimentos realizados para evaluar genotipos depende importantemente del grado de correspondencia entre la situación experimental real y el modelo utilizado para representarla. En este estudio se presentan las esperanzas de cuadrados medios de cuatro modelos lineales, cada uno con tres variantes, utilizados en la evaluación de genotipos en  $l$  localidades durante  $a$  años en forma tal que el fitomejorador puede identificar la forma, para cada modelo, de probar las hipótesis de interés. Adicionalmente, se discute la naturaleza de los factores y las relaciones entre ellos, todo esto en el contexto de situaciones experimentales típicas en la evaluación de genotipos correspondientes a cultivos anuales.

### PALABRAS CLAVE ADICIONALES

Modelo sin restricciones, interacción genético-ambiental, esperanzas de cuadrados medios, modelo aleatorio, modelo mixto.

### SUMMARY

The quality of the analysis of the data produced in combined experiments conducted to evaluate genotypes depends primarily on the degree of correspondence between the real experimental situation and the model used to represent it. In this study, the expected mean squares of each of four models, each including three cases, used to represent situations in which "g" genotypes are evaluated in "l" locations during a years are presented: thus plant breeders can identify for each case the procedure to test hypotheses. Additionally, the nature and relationships among factors are discussed in the

context of the typical experimental situations relative to genotypic evaluation of annual crops.

### ADDITIONAL INDEX WORDS

Model without restrictions, genotype x environment interaction, expected mean squares, random model, mixed model.

### INTRODUCCION

Debido al problema que representa la interacción entre genotipos y ambientes, el fitomejorador siempre se ve en la necesidad de planear la evaluación de genotipos a través de varias localidades durante dos o más años. Una vez realizados los experimentos, la fase siguiente es el análisis de la información producida, seguida de su interpretación. En la fase de análisis usualmente se incluye la realización de análisis de varianza combinados con datos producidos en los experimentos en cuestión. Ciertamente, el análisis de varianza (ANAVA) no produce información que signifique la solución a todas las interrogantes que surjan en la evaluación genotípica. Sin embargo, su utilidad ha sido muy importante en el fitomejoramiento en general y en el estudio de la interacción entre genotipos y ambientes en particular. Sahagún (1990) señala que además de ser una herramienta para detectar la significancia de la interacción, el ANAVA ha sido exitosamente utilizado para diversos propósitos relacionados con el problema de la interacción.

<sup>1</sup> Depto. de Fitotecnia, UACH; Chapingo, Méx. C.P. 56230.

Por ejemplo, el ANAVA ha sido útil para la identificación de áreas homogéneas desde el punto de vista de que no se produce interacción significativa. Este uso es importante pues si bien el fitomejorador prácticamente no tiene la posibilidad de predecir las características agrónomicamente importantes asociadas con los años (cantidad y distribución de lluvias, heladas, granizadas, etc.), sí cuenta con la relativa estabilidad de las características más distintivas de las localidades (tipo de suelo, altura sobre el nivel del mar, latitud, etc.). De esta manera se hace posible contar con un criterio para seleccionar las localidades para efectos de evaluación y para identificar localidades que producen información redundante. Numerosos estudios han sido desarrollados en esta temática (Horner y Frey, 1957; McCain y Schultz, 1959; Liang *et al.*, 1966; Ghaderi *et al.*, 1980).

La evaluación de la estabilidad de los genotipos también ha sido realizada con base en información generada por el ANAVA (Plaisted y Peterson, 1959; Lin y Binns, 1988). Entre otros usos del análisis de varianza se incluyen: 1) El estudio de las propiedades de grupos de genotipos formados con base en su estructura genotípica o en sus patrones de respuesta (Eberhart y Russell, 1969; Weatherspoon, 1970), y 2) la producción de información útil en el desarrollo de estrategias de uso óptimo de recursos y de maximización de ganancias genéticas en programas de fitomejoramiento (Sprague y Federer, 1951; Barker *et al.*, 1981).

Si bien los usos anteriores constituyen una muestra que ilustra perfectamente el potencial del ANAVA, éste puede producir resultados erróneos si ciertas condiciones no se satisfacen. Entre las condiciones que deben satisfacerse, de acuerdo con los análisis típicos, se encuentran las relativas a la distribución normal de la variable en estudio; la homogeneidad de varianzas; la

independencia de errores y la relación funcional correcta. Es con esta última cuestión con la que se relaciona más estrechamente el propósito de esta presentación.

La relación funcional, desde luego, tiene que ver con el modelo que se utiliza para representar la situación experimental con base en el cual se desarrolla el análisis de varianza. Evidentemente, debe haber una estrecha relación entre el modelo utilizado y la calidad de la información derivada de su análisis.

Es común que cuando se va a realizar un análisis combinado no se reflexione sobre cuál deberá ser un modelo adecuado en que se base dicho análisis. Sahagún (1992) indica que, frecuentemente el modelo utilizado, consciente o inconscientemente, se toma de los que se encuentran al alcance del investigador en la literatura sobre el tema, pudiendo ocurrir que no se haga una selección satisfactoria. Adicionalmente, si bien es cierto que el uso de paquetes estadísticos permite la determinación de las fuentes de variación por parte del usuario, no ofrece la posibilidad de especificar algunos aspectos que son decisivos en la determinación de las esperanzas de cuadrados medios que, a fin de cuentas, son las que determinan las formas de realizar las pruebas de hipótesis y la estimación de componentes de varianza. Por ejemplo, los modelos pueden incluir o no restricciones sobre los parámetros pudiendo ser éste un factor que afecte la forma de realizar las pruebas de hipótesis. Como indica Sahagún (1992), este tipo de casos son el reflejo de que el tema de la delimitación y selección de modelos para series de experimentos no ha sido debidamente considerado.

En esta presentación se discuten cuatro modelos lineales para la representación de situaciones experimentales en que se evalúan  $g$  genotipos en series de experimentos (considerando, preferentemente, " $l$ " localidades

y "a" años). En este trabajo también se estudian tanto las relaciones entre los factores como su naturaleza. Los análisis de varianza se presentan de manera tal que se puedan determinar los efectos de la ausencia de restricciones en las pruebas de hipótesis correspondientes a cada una de las fuentes de variación.

## TEORIA Y METODOLOGIA

En las series típicas de experimentos para evaluar genotipos se tiene en consideración un número de  $r \times a \times \ell \times g$  parcelas experimentales para evaluar "g" genotipos durante "a" años en cada una de " $\ell$ " localidades utilizando en cada combinación de localidad y año un diseño en, digamos, bloques al azar con  $r$  repeticiones. Salvo excepciones, explícitamente referidas, estas consideraciones serán el punto de partida de este estudio.

### Factores de Clasificación y Modelos

Aunque no con las tres variantes, tres de los cuatro modelos a estudiarse ya han sido descritos en un trabajo previo (Sahagún, 1992); sin embargo, por ser parte central de este estudio a continuación se hace una presentación de todos ellos. En series de experimentos domina la consideración de que los factores "años" (A), "localidades" (L) y "genotipos" (G) son factores cruzados (Sprague y Federer, 1951; Comstock y Moll, 1963). Al factor "bloques" o repeticiones (B) es común que se le considere anidado en localidades (B/L). Con estas consideraciones como base (los factores A, L y G cruzados y el factor B anidado en L) se ha generado un modelo lineal para explicar el comportamiento del genotipo  $i$  del bloque  $j$  de la localidad  $k$  en el año  $m$  que, omitiendo detalles obvios, es de la forma

$$Y_{ijkm} = \mu + G_i + (B/L)_{jk} + L_k + (GL)_{ik} + A_m + (GA)_{im} + (LA)_{km} + (GLA)_{ikm} + E_{ijkm} \quad (1)$$

Otros autores, Campbell y Lafever (1977), Brennan y Byth (1979) y McIntosh (1983), entre otros, por su parte, han utilizado un modelo de la forma

$$Y_{ijkm} = \mu + G_i + (B/AL)_{jkm} + L_k + (GL)_{ik} + A_m + (GA)_{im} + (LA)_{km} + (GLA)_{ikm} + E_{ijkm} \quad (2)$$

Con respecto al modelo (1), lo distintivo del modelo (2) es que el factor B se considera anidado tanto en A como en L. Esta consideración se puede referir a una situación en la que cada bloque ocurre una sola vez (no es el mismo) en cada combinación de año y localidad. Si los factores A y L son auténticamente cruzados, la consideración de anidamiento de B en A y L es correcta cuando los bloques de cada localidad son diferentes en cada año. Ahora bien, si el factor B se considera anidado sólo en el L cuando A y L son factores cruzados (se enfatiza) lo que se representa es una situación experimental en que en cada localidad, estrictamente hablando, los bloques que se utilizan son los mismos cada año. La diferencia entre los modelos (1) y (2) no es trivial, tiene implicaciones en la estructura del error que se reflejan en que, por ejemplo, la prueba de hipótesis para la interacción GAL se realiza de manera diferente como se verá más adelante.

Considerando que las componentes ambientales asociadas con los años, dado un año particular, no se presentan de igual manera a través de las localidades y que las componentes ambientales básicas asociadas con las localidades se mantienen esencialmente constantes de un año a otro, resulta razonable considerar la existencia de una relación de anidamiento: A anidado en L. Con esta consideración, suponiendo que los bloques son diferentes de un año a otro, el factor B debería estar anidado en el A. Para estas condiciones, Sahagún (1990) consideró un modelo de la forma

$$Y_{ijkm} = \mu + G_i + (B/A/L)_{jkm} + (A/L)_{km} + (G(A/L))_{ikm} + L_m + (GL)_{im} + E_{ijkm} \quad (3)$$

Ahora bien, si las componentes ambientales asociadas tanto con los años como con las localidades se manifestaran de manera diferente en tiempo y en espacio, como en el caso en que en cada año se utilicen localidades diferentes, quizá para ganar en el tamaño de muestra de localidades, la situación experimental demandaría de un modelo diferente. En este caso los factores A y L estarían confundidos. Así, las axl combinaciones de estos factores confundidos podrían ser consideradas como los niveles de otro factor, digamos C. Para esta situación experimental se podría recurrir a la expresión

$$Y_{ijp} = \mu + G_i + (B/C)_{jp} + C_p + (GC)_{ip} + E_{ijp} \quad (4)$$

para explicar el comportamiento del genotipo i en el bloque j del ambiente (combinación de año y localidad) p. Muchos autores han estudiado la interacción entre genotipos y ambientes utilizando como ambientes a cada una de las combinaciones de localidades y años (Eberhart y Russell, 1966).

### Tipos de Factores

Con el fin de abordar exitosamente el tema de las pruebas de hipótesis es fundamental definir para cada factor si éste es fijo o aleatorio. El factor G (genotipos) para el caso de evaluaciones con el fin de hacer selección deberá considerarse como un factor de efectos fijos. Sin embargo, si las evaluaciones son realizadas con el propósito de estimar componentes de varianzas genéticas, como cuando se desea estudiar la estructura de variabilidad genética de una población, el factor genotipos (o grupos de genotipos) deberá ser considerado como aleatorio. Para el primer caso los genotipos

en evaluación son los que interesa comparar, en tanto que para el estudio de la estructura de variabilidad genética, los genotipos en evaluación constituyen una muestra aleatoria de la población en estudio.

Con respecto al factor A (años), ciertamente es difícil concebir que los años, considerados por su número cronológico, puedan ser conceptuados como una muestra aleatoria. Sin embargo, resulta incuestionable que dicho factor produce efectos que son impredecibles en el comportamiento de un cultivo. Para el caso del factor B (bloques) la situación es completamente análoga. Todo esto es particularmente cierto en el caso de cultivos anuales en condiciones de campo. En el otro extremo, digamos agricultura en condiciones de invernadero con control de temperatura, humedad, etc. el factor "años" puede tener poco significado como fuente de variación.

Con referencia al factor L (localidades), desde el punto de vista de los objetivos de la evaluación genotípica, éste debería ser de naturaleza aleatoria. Evidentemente, en la vida real los campos experimentales comúnmente se tienen definidos *a priori* y en ocasiones no son congruentes con los objetivos de la evaluación, ya que su elección con frecuencia obedece a cuestiones ajenas a éstos (donaciones, expropiaciones, etc.). Aún en estas circunstancias se podría hacer una selección aleatoria de localidades (con ciertas limitaciones de aplicación por tratarse de un número finito de las que son elegibles). Es en este sentido que Ford y Nielsen (1982) desarrollaron un sistema computarizado para delimitar áreas a las cuáles pudieran ser extrapolados los resultados obtenidos en los campos experimentales, considerando aspectos tales como tipo de suelo: fecha promedio de ocurrencia de primera y última helada: precipitación promedio anual, etc. Sin embargo, es de consignarse que ya es común que las pruebas del material genético

se hagan también en campos de agricultores cooperantes, lo cual hace más representativo al factor localidades.

El tratamiento que se le dará al factor localidades en esta presentación será el correspondiente al de un factor de efectos aleatorios que sigue una distribución normal con media cero y una cierta varianza finita,  $V(L)$  digamos. El caso en que se trata de una población finita de niveles lo mencionan varios autores (Cornfield y Tukey, 1956; Searle y Fawcet, 1970; Searle, 1971).

### Consideraciones Adicionales

No se hará referencia a todo lo que se pudiera considerar en la construcción de un modelo. Más bien se pretende ultimar los detalles necesarios para tener una definición completa de los modelos (1), (2), (3) y (4). Sólo se hará mención a las propiedades distribucionales de los términos aleatorios y a las restricciones, en su caso, que se suelen imponer sobre algunos de los parámetros de los modelos. Con ésto como base se podrán derivar las formas de realizar las pruebas de hipótesis.

Para el caso de factores, o interacciones, de efectos aleatorios, éstos serán considerados como variables aleatorias normal e independientemente distribuídas con media cero y una varianza finita, constante para cada uno de los elementos del conjunto de niveles de un mismo factor, interacción o cualquier tipo de partición de tipo aleatorio.

Ahora, con relación a las restricciones sobre parámetros del modelo, es frecuente que para los niveles de un factor de efectos fijos, digamos  $G$ , se establezcan en forma explícita en el modelo expresiones de la forma

$$\sum_{i=1}^g G_i = 0$$

De la misma manera, para una interacción que involucre un factor de efectos fijos y al menos uno de efectos aleatorios, digamos  $(GA)_{ij}$ , es común que se imponga la restricción de que la suma sobre todos los niveles del factor de efectos fijos sea cero; es decir, por ejemplo, si  $G$  es de efectos fijos y  $A$  es de efectos aleatorios, la restricción impuesta, en su caso, es:

$$\sum_{i=1}^g G_{ij} = 0 \text{ para toda } j$$

### Esperanzas de Cuadrados Medios

Con respecto a la determinación de las esperanzas de cuadrados medios, una vez definido el modelo (sin ser necesario establecer la distribución que sigan los términos aleatorios), ésta se puede realizar con base en el uso de un resultado que asociado al modelo lineal general, expresado éste en la forma

$$\underline{Y} = X \underline{B} + \underline{E}' \quad , \quad E(\underline{E}') = \underline{0} \quad \text{y} \\ V(\underline{E}') = V,$$

establece que

$$E(\underline{Y}'A\underline{Y}) = \text{Tr}(AV) + (E(\underline{Y}))'A(E(\underline{Y}))$$

en donde  $E$ ,  $V$  y  $\text{Tr}$  denotan "esperanza matemática", "varianza" y "traza", respectivamente. Una forma alternativa, pero más práctica, para determinar estas esperanzas consiste en la aplicación de reglas desarrolladas para tal propósito para modelos con restricciones (John, 1971) y sin restricciones (Searle, 1971).

### Prueba de Hipótesis

Las pruebas de hipótesis se considerarán con base en el concepto de modelo insesgado. Este establece que un modelo es insesgado para la prueba de una hipótesis si existen dos cuadrados medios cuyas esperanzas sean idénticas sí y sólo sí la hipótesis a probar

es cierta. Cuando se tiene insesgo para una hipótesis, resulta que la prueba es exacta. Si no existe la situación de insesgo referida se apelará al recurso del procedimiento establecido por Satterthwaite (1946).

## RESULTADOS

**Modelo (1).** En este modelo se considera que los factores A, L y G se encuentran cruzados y que B está anidado sólo en L. Con respecto a los casos con restricciones (modelo mixto, con G como único factor de efectos fijos) y sin restricciones (modelo mixto y modelo aleatorio) las formas de realizar las pruebas de hipótesis difieren sólo para los términos A, L y AL (Cuadro 1).

**Modelo (2).** La diferencia de éste con respecto al modelo (1) es que en el modelo (2) el factor B se considera anidado tanto en A como en L y en el modelo (1) el anidamiento de B es sólo en el factor L. En el ANAVA, esta diferencia produce que el error correspondiente al modelo (1) tenga una suma de cuadrados que, con relación a la del error del modelo (2), contiene adicionalmente la suma de cuadrados debida a otra fuente de variación. Por supuesto, también se presenta otra diferencia en otra fuente de variación, la correspondiente al factor B (Cuadro 2).

Estas diferencias implican, a su vez, cambios en la realización de algunas pruebas de hipótesis. Por ejemplo, en el modelo (1) con restricciones la prueba de L se realiza con el cociente cuyo numerador es la suma de dos cuadrados medios, el de L y el del error y cuyo denominador es también la suma de dos cuadrados medios, el de AL y el de B/L. En el modelo (2), en cambio, la misma prueba se realiza con el cociente cuyo numerador y denominador son el cuadrado medio de L y el de A/L, respectivamente.

Con respecto a las fuentes de variación en que está involucrado el factor G, sólo en el

caso de GAL se presentan diferencias cuando se consideran los modelos (1) y (2). Si bien en ambos modelos se utiliza el cociente  $C_g/C_g$ ,  $C_g$  (el cuadrado medio del error) tiene una estructura que difiere de un modelo a otro. Además, con el modelo (2) las pruebas de hipótesis para A, L y AL resultan ser aproximadas cuando no se consideran las restricciones, mientras que cuando éstas sí se incluyen las pruebas de hipótesis resultan ser exactas.

**Modelo (3).** En este modelo el factor B se encuentra anidado en A y éste en L. En el caso en que se establecen restricciones sobre los parámetros cuando sólo el factor G es fijo, todas las pruebas de hipótesis resultan ser exactas (Cuadro 3). Sin embargo, cuando no se consideran las restricciones las pruebas para los factores L y A/L son sólo aproximadas.

Ahora bien, con respecto a los factores G, GL, y G(A/L) en este modelo, con o sin restricciones, las pruebas resultan ser exactas e iguales. También en los modelos (1) y (2) las pruebas para fuentes de variación en que está involucrado el factor G son de la misma forma, con restricciones o sin ellas.

**Modelo (4).** En este modelo, caracterizado por la confusión de los factores A y L, la única prueba aproximada es la correspondiente al factor C (ambientes) cuando en el modelo no se incluyen restricciones sobre los parámetros (Cuadro 4). Las pruebas de hipótesis, para G y GC, importantes desde el punto de vista del fijomejorador, resultan ser pruebas exactas con o sin restricciones. La única prueba aproximada es la prueba para ambientes cuando no se usan restricciones.

Es de consignarse que algunos de los resultados mostrados en este capítulo ya han sido presentados en otros estudios. Por ejemplo, McIntosh (1983) hace referencia al modelo (2), pero de éste no presenta las

Cuadro 1. Esperanzas de cuadrados medios y pruebas de hipótesis correspondientes al modelo (1).

Fuente de variación	Esperanzas de cuadrados medios <sup>1</sup>										Pruebas de F	
	CM	E	GAL	GL	GA	G	B/AL	AL	L	A	SR <sup>2</sup>	CR <sup>3</sup>
Años (A)	C <sub>1</sub>	p	p <sup>4</sup>		p <sup>4</sup>			p		p	(C <sub>1</sub> +C <sub>8</sub> )/(C <sub>3</sub> +C <sub>6</sub> )	C <sub>1</sub> /C <sub>3</sub>
Localidades (L)	C <sub>2</sub>	p	p <sup>4</sup>	p <sup>4</sup>			p	p	p		(C <sub>2</sub> +C <sub>8</sub> +C <sub>9</sub> )/(C <sub>3</sub> +C <sub>4</sub> +C <sub>7</sub> )	(C <sub>2</sub> +C <sub>9</sub> )/(C <sub>3</sub> +C <sub>4</sub> )
AL	C <sub>3</sub>	p	p <sup>4</sup>					p			C <sub>3</sub> /C <sub>8</sub>	C <sub>3</sub> /C <sub>4</sub>
Bloques (B)/L	C <sub>4</sub>	p					p				C <sub>4</sub> /C <sub>9</sub>	C <sub>4</sub> /C <sub>9</sub>
Genotipos (G)	C <sub>5</sub>	p	p	p	p	p					(C <sub>5</sub> +C <sub>8</sub> )/(C <sub>6</sub> +C <sub>7</sub> )	(C <sub>5</sub> +C <sub>8</sub> )/(C <sub>6</sub> +C <sub>7</sub> )
GA	C <sub>6</sub>	p	p		p						C <sub>6</sub> /C <sub>8</sub>	C <sub>6</sub> /C <sub>8</sub>
GL	C <sub>7</sub>	p	p	p							C <sub>7</sub> /C <sub>8</sub>	C <sub>7</sub> /C <sub>8</sub>
GAL	C <sub>8</sub>	p	p								C <sub>8</sub> /C <sub>9</sub>	C <sub>8</sub> /C <sub>9</sub>
Error (E)	C <sub>9</sub>	p										

<sup>1</sup> Varianzas o pseudovarianzas de los factores indicados. <sup>2</sup> SR= Modelo aleatorio o mixto (sólo G es fijo) sin restricciones.

<sup>3</sup> CR= Sólo G es fijo. Modelo con restricciones. <sup>4</sup> p = Presencia del término; p<sup>4</sup> = Presencia sólo en modelos aleatorio y mixto sin restricciones.

Cuadro 2. Esperanzas de cuadrados medios y cocientes para probar hipótesis (Modelo 2).

Fuente de variación	Esperanzas de cuadrados medios <sup>1</sup>										Pruebas de F	
	CM	E	GAL	GL	GA	G	B/AL	AL	L	A	SR <sup>2</sup>	CR <sup>3</sup>
Años (A)	C <sub>1</sub>	p	p <sup>4</sup>		p <sup>4</sup>		p	p		p	(C <sub>1</sub> +C <sub>8</sub> )/(C <sub>3</sub> +C <sub>6</sub> )	C <sub>1</sub> /C <sub>3</sub>
Localidades (L)	C <sub>2</sub>	p	p <sup>4</sup>	p <sup>4</sup>			p	p	p		(C <sub>2</sub> +C <sub>8</sub> )/(C <sub>3</sub> +C <sub>7</sub> )	C <sub>2</sub> /C <sub>3</sub>
AL	C <sub>3</sub>	p	p <sup>4</sup>				p	p			(C <sub>3</sub> +C <sub>9</sub> )/(C <sub>4</sub> +C <sub>8</sub> )	C <sub>3</sub> /C <sub>4</sub>
Bloques (B)/L	C <sub>4</sub>	p					p				C <sub>4</sub> /C <sub>9</sub>	C <sub>4</sub> /C <sub>9</sub>
Genotipos (G)	C <sub>5</sub>	p	p	p	p	p					(C <sub>5</sub> +C <sub>8</sub> )/(C <sub>6</sub> +C <sub>7</sub> )	(C <sub>5</sub> +C <sub>8</sub> )/(C <sub>6</sub> +C <sub>7</sub> )
GA	C <sub>6</sub>	p	p		p						C <sub>6</sub> /C <sub>8</sub>	C <sub>6</sub> /C <sub>8</sub>
GL	C <sub>7</sub>	p	p	p							C <sub>7</sub> /C <sub>8</sub>	C <sub>7</sub> /C <sub>8</sub>
GAL	C <sub>8</sub>	p	p								C <sub>8</sub> /C <sub>9</sub>	C <sub>8</sub> /C <sub>9</sub>
Error (E)	C <sub>9</sub>	p										

<sup>1</sup> Varianzas o pseudovarianzas de los factores indicados. <sup>2</sup> SR= Modelo aleatorio o mixto (sólo G es fijo) sin restricciones.

<sup>3</sup> CR= Modelo con restricciones, sólo G fijo. <sup>4</sup> p = Presencia del término; p<sup>4</sup> = Presencia sólo en modelos aleatorio y mixto sin restricciones.

Cuadro 3. Esperanzas de cuadrados medios y pruebas de hipótesis correspondientes al modelo 3.

Fuente de variación	Cuadrados medios	Esperanzas de cuadrados medios <sup>1</sup>							Pruebas de F	
		E	G(A/L)	GL	G	B/A/L	A/L	L	SR <sup>2</sup>	CR <sup>3</sup>
Localidades (L)	C <sub>1</sub>	p	p <sup>4</sup>	p		p	p	p	(C <sub>1</sub> +C <sub>6</sub> )/(C <sub>2</sub> +C <sub>5</sub> )	C <sub>1</sub> /C <sub>2</sub>
Años (A)/L	C <sub>2</sub>	p	p <sup>4</sup>			p	p		(C <sub>2</sub> +C <sub>7</sub> )/(C <sub>3</sub> +C <sub>6</sub> )	C <sub>2</sub> /C <sub>3</sub>
Bloques (B)/A/L	C <sub>3</sub>	p				p			C <sub>3</sub> /C <sub>7</sub>	C <sub>3</sub> /C <sub>7</sub>
Genotipos (G)	C <sub>4</sub>	p	p	p	p				C <sub>4</sub> /C <sub>5</sub>	C <sub>4</sub> /C <sub>5</sub>
GL	C <sub>5</sub>	p	p	p					C <sub>5</sub> /C <sub>6</sub>	C <sub>5</sub> /C <sub>6</sub>
G (A/L)	C <sub>6</sub>	p	p						C <sub>6</sub> /C <sub>7</sub>	C <sub>6</sub> /C <sub>7</sub>
Error (E)	C <sub>7</sub>	p								

<sup>1</sup> Varianzas o pseudovarianzas de los factores indicados. <sup>2</sup> SR= Modelo aleatorio o mixto (sólo G es fijo) sin restricciones. <sup>3</sup> CR= Sólo G es fijo. Modelo con restricciones. <sup>4</sup> p = Presencia del término; p<sup>4</sup> = Presencia sólo en el modelo aleatorio y en el mixto sin restricciones.

Cuadro 4. Esperanzas de cuadrados medios y pruebas de hipótesis correspondientes al modelo 4.

Fuente de Variación	Cuadrados medios	Esperanzas de cuadrados medios <sup>1</sup>					Pruebas de F	
		E	GC	G	B/C	C	SR <sup>2</sup>	CR <sup>3</sup>
Ambientes (C)	C <sub>1</sub>	p	p <sup>4</sup>		p	p	(C <sub>1</sub> +C <sub>3</sub> )/(C <sub>2</sub> +C <sub>4</sub> )	C <sub>1</sub> /C <sub>2</sub>
Bloques (B)/C	C <sub>2</sub>	p			p		C <sub>2</sub> /C <sub>5</sub>	C <sub>2</sub> /C <sub>5</sub>
Genotipos (G)	C <sub>3</sub>	p	p	p			C <sub>3</sub> /C <sub>4</sub>	C <sub>3</sub> /C <sub>4</sub>
GC	C <sub>4</sub>	p	p				C <sub>4</sub> /C <sub>5</sub>	C <sub>4</sub> /C <sub>5</sub>
Error	C <sub>5</sub>	p						

<sup>1</sup> Varianzas o pseudovarianzas de los factores indicados. <sup>2</sup> SR= Modelo aleatorio o mixto (sólo G es fijo) sin restricciones. <sup>3</sup> CR= Sólo G es fijo. Modelo con restricciones. <sup>4</sup> p = Presencia del término; p<sup>4</sup> = Presencia sólo en el modelo aleatorio y en el mixto sin restricciones.

esperanzas de cuadrados medios ni el caso del modelo mixto sin restricciones. Por su parte, Sahagún (1992) estudia los modelos (1), (2) y (3) pero en éstos no incluye el caso de modelo mixto sin restricciones en los parámetros.

## DISCUSION

Ya se ha enfatizado que el ANAVA por sí solo no es capaz de dar respuesta a muchas de las interrogantes que se pueden generar con relación a la interacción entre

genotipos y ambientes. Inclusive, las respuestas que se pueden generar tienen un valor que en parte depende de cuestiones no imputables al ANAVA como tal. La forma en que se modela una situación experimental es de una gran importancia en la calidad de la información obtenida. Por la naturaleza de los conceptos utilizados en la delineación de los modelos, éstos con frecuencia llegan a adolecer de importantes discrepancias con respecto a la realidad. Por ejemplo, si bien los conceptos de anidamiento, cruzamiento, etc. son claros, existen otros que con mucha

frecuencia son difíciles de establecer con precisión.

En realidad, la definición de los conceptos de localidad y año, en el contexto que nos ocupa, no siempre es fácil de precisar de una manera eficiente. ¿Cuáles son los elementos definitorios de una localidad?, ¿es el tiempo meteorológico parte de una localidad?, ¿son los niveles tecnológico, económico, cultural, etc. parte de una localidad? Sin duda, la búsqueda de la respuesta a estas interrogantes, y a otras similares, arrojará información que será muy útil en el estudio del problema de la interacción entre genotipos y ambientes.

En los modelos (1) y (2) se considera que los factores L y A son cruzados. Como lo señala Sahagún (1992), la parte menos realista de esta suposición es la que implica que en cada año las manifestaciones de los factores ambientales asociados con los años son iguales en las diferentes localidades; en ocasiones las diferencias llegan a ser verdaderamente sustanciales, produciendo un impacto diferencial muy importante, particularmente en condiciones de ausencia de irrigación y heterogeneidad ambiental dentro de la región que es objeto de un programa de mejoramiento, situación común en la agricultura de cultivos anuales de temporal.

El modelo (3), al parecer, representa de una manera más aproximada una situación que se presenta en la evaluación de genotipos en diferentes lugares y años en el sentido de que, si bien no es una verdad absoluta, sí se reconoce que en un medio ambiental natural, como el prevaleciente en muchas regiones donde se practica agricultura de cultivos anuales en condiciones de temporal, el factor "años", en realidad, no está formado por niveles identificados por los años cronológicos. Durante un año, cada localidad recibe manifestaciones particulares de los aspectos ambientales considerados como inherentes al factor "años". Por otra

parte, si resultara razonable suponer que los componentes del ambiente más directamente asociados con el factor "localidades" se mantienen, dentro de cada localidad, invariantes a través de los años, entonces el factor "años" debe considerarse anidado en el factor "localidades".

Debe ser evidente que en el modelo (4) se presenta una situación particular con respecto al concepto de balance: El problema de preservar, para una mayor facilidad de análisis, el balance con respecto a los factores A y L, es obvia. Además, y ya en el contexto de la evaluación genotípica, el considerar el ambiente como un factor en el que los niveles son cada combinación de año y localidad, tiene un significado importante desde el punto de vista de interpretación puesto que son los que constituyen el ambiente natural en el que normalmente se desarrollan los cultivos agrícolas anuales. Sin embargo, si bien tiene la ventaja de ser conceptual y metodológicamente más sencillo, en este enfoque no se tiene la posibilidad de generar información relativa a los aspectos constantes del ambiente asociados a las localidades. Esto, sin embargo, también se compensa por el estudio de la estabilidad del comportamiento de los genotipos a través de los ambientes.

En lo referente a los tipos de factores, que pueden ser fijos o aleatorios, por lo que respecta a los factores A y L en este estudio siempre fueron considerados como aleatorios. Se procedió de esta manera tratando de presentar la situación más común con respecto a los objetivos del mejoramiento en el sentido de que las inferencias sobre el comportamiento de los genotipos no aludan exclusivamente a los años y localidades utilizados en la evaluación, sino a una población de años y localidades de donde los considerados son una muestra al azar. Esto, obviamente, puede representar dificultades para su concreción en la realidad. No obstante, si por alguna razón especial se

deseara comparar los efectos de, digamos, localidades, éstas se podrían considerar como de efectos hijos.

El factor bloque (B) se incluyó porque normalmente permite el control de proporciones importantes de variación local, debida, en experimentos de campo, a la heterogeneidad del suelo principalmente (desde luego, diseños experimentales diferentes al de bloques al azar también pueden ser utilizados). La consideración del factor bloques como factor de efectos aleatorios también obedece a cuestiones de tipo aplicado; normalmente no se consideran procedimientos aleatorios para su selección.

Por otra parte, con relación a los cuatro modelos es notable que en todas las fuentes de variación en que está involucrado el factor G las esperanzas de cuadrados medios para cada modelo son las mismas, independientemente de si se incluyen o no las restricciones sobre los parámetros. Esto significa que tanto la estimación de componentes de varianza (en su caso) como la realización de pruebas de hipótesis de estas fuentes de variación no dependen de la imposición o no de restricciones.

Por el contrario, salvo en el modelo (1), en las pruebas de hipótesis de las fuentes de variación cuyas esperanzas de cuadrados medios discrepan, el uso de restricciones implica pruebas de hipótesis exactas en tanto que en el modelo sin restricciones dichas pruebas son aproximadas. Estas diferencias, por supuesto, también tienen que ver con la forma en que el modelo se construye; las restricciones, o su ausencia, también deben ser un reflejo de la situación experimental: las esperanzas de los cuadrados medios dependen fundamentalmente de la forma en que expliquemos a las observaciones.

## BIBLIOGRAFIA

- Barker, R.E., A.W. Hoving, I.T. Carlson, P.N. Droslom, D.A. Sleper, J.G. Ross, and M.D. Casler. 1981. Genotype-environment interactions for forage yield of red canary grass clones. *Crop Sci.* 21:567-571.
- Brennan, P.S., and D.E. Byth. 1979. Genotype x environmental interactions for wheat yields and selection for widely adapted wheat genotypes. *Aust. J. Agric. Res.* 30:221-232.
- Campbell, L.G., and H.N. Lafever. 1977. Cultivar x environmental interactions in soft red winter wheat yield test. *Crop Sci.* 17:604-608.
- Comstock, R.E., and R.H. Moll. 1963. Genotype x environment interactions. Symposium on statistical genetics and plant breeding. NAS-CRC. pp 164-196.
- Cornfield, J., and J.W. Tukey. 1956. Average values of means squares in factorials. *Ann. Math. Statist.* 27:907-949.
- Eberhart, S.A., and W.A. Russell. 1966. Stability parameters for comparing varieties. *Crop Sci.* 6:36-40.
- \_\_\_\_\_. 1969. Yield and stability for a 10-line diallel of single-cross and double-cross maize hybrids. *Crop Sci.* 9:357-361.
- Ford, S.L., and G.A. Nielsen. 1982. Selecting experimentally analogous areas for agronomic research using computer graphics. *Agron. J.* 74:261-265.
- Ghaderi, A., E.H. Everson, and C.E. Cross. 1980. Classifications of environments and genotypes in wheat. *Crop Sci.* 20:707-710.
- Horner, T.W., and K.J. Frey. 1957. Methods for determining natural areas for oat varietal recommendations. *Agron. J.* 49:313-315.
- John, P.W.M. 1971. *Statistical Design and Analysis of Experiments.* The Macmillan Company, New York.

- Liang, G.H.L., E.G. Heyne, and T.L. Walter. 1966.** Estimates of variety-environment interactions in yield test of three small grains and their significance on the breeding programs. *Crop Sci.* 6:135-139.
- Lin, C.S., and M.R. Binns. 1988.** A method of analyzing cultivar x location x year experiments: a new stability parameter. *Theor. Appl. Genet.* 76:425-430.
- McCain, F.S., and E.F. Schultz, Jr. 1959.** A method for determining areas for corn varietal recommendations. *Agron. J.* 51:476-478.
- McIntosh, M.S. 1983.** Analysis of combined experiments. *Agron. J.* 75:153-155.
- Plaisted, R.L., and Peterson, L.C. 1959.** A technique for evaluating the ability of selections to yield consistently in different locations and seasons. *Am. Potato J.* 36:381-385.
- Sahagún C., J. 1990.** La utilidad del análisis de varianza en el estudio de la interacción entre genotipos y ambientes. *Xilonen* 1:21-32.
- \_\_\_\_\_. 1992. Consideraciones preliminares sobre la evaluación genotípica en series de experimentos conducidos en invernadero. *Revista Chapingo* 79-80:26-31.
- Satterthwaite, F.R. 1946.** An approximate distribution of estimates of variance components. *Biometrics Bull.* 2:110-112.
- Searle, S.R. 1971.** *Linear Models.* Wiley, New York.
- \_\_\_\_\_, and R.F. Fawcett. 1970. Expected mean squares in variance component models having finite populations. *Biometrics* 26:243-254.
- Sprague, F.S., and W.T. Federer. 1951.** A comparison of variance components in corn yield trials. II. Error, year x variety, location x variety and variety components. *Agron. J.* 43:531-541.
- Weatherspoon, J.H. 1970.** Comparative yields of single, three-way, and double crosses of maize. *Crop Sci.* 10:157-159.